



Text Mining and Sentiment Analysis for Strategic Market Intelligence

Dr. Jagadish Gurrula

Department of CSE, Koneru Lakshmaiah Education Foundation Green Fields, Guntur, Andhra Pradesh, India
gjagadish@kluniversity.in

ABSTRACT: This study leverages text mining and sentiment analysis techniques to extract actionable insights from unstructured textual data for strategic market intelligence in competitive environments.

KEYWORDS: Text Mining, Sentiment Analysis, Market Intelligence, Natural Language Processing, Social Media Analytics, Opinion Mining,

I. INTRODUCTION

In the age of digital transformation, enterprises are inundated with vast volumes of unstructured textual data from sources like social media, customer reviews, news articles, and forums. Traditional market research techniques, while valuable, often lack the agility and depth required to process such data in real time. Text mining and sentiment analysis offer powerful tools to address this challenge by extracting meaningful patterns and emotional tones from textual content.

Strategic market intelligence involves the collection and analysis of data to inform decision-making related to competitive positioning, customer preferences, and emerging trends. By applying natural language processing (NLP) and machine learning algorithms, organizations can transform raw text into structured insights, thereby enabling timely and data-driven strategies. This paper explores how text mining and sentiment analysis contribute to market intelligence, highlighting applications, techniques, and outcomes that enhance strategic business capabilities.

II. LITERATURE REVIEW

The role of text mining in business analytics has grown significantly in recent years, with studies underscoring its utility in customer experience management, brand monitoring, and product development. Feldman and Sanger (2007) defined text mining as the process of deriving high-quality information from text, involving tasks such as information retrieval, pattern recognition, and linguistic analysis.

Sentiment analysis, a subfield of text mining, classifies the emotional polarity of text—positive, negative, or neutral. Liu (2012) emphasized its significance in consumer behavior analysis and reputation management. Tools like VADER, TextBlob, and machine learning classifiers (SVM, Naïve Bayes, deep learning) have been widely adopted in sentiment classification tasks.

Recent works have applied these methods to various industries: Chevalier and Mayzlin (2006) studied online book reviews to analyze competitive dynamics; Pang and Lee (2008) investigated movie reviews; and broader applications now span sectors like retail, finance, and healthcare.

III. RESEARCH METHODOLOGY

1. Data Collection:

Text data was sourced from:

- Twitter (via APIs) – mentions of select products and brands
- Amazon product reviews
- Business news headlines



2. Preprocessing:

Data underwent tokenization, stop-word removal, lemmatization, and noise filtering (e.g., hashtags, mentions, special characters).

3. Text Mining Techniques:

- **TF-IDF:** Used for keyword extraction and feature weighting
- **Topic Modeling (LDA):** Identified recurring themes in product reviews and social posts
- **Named Entity Recognition (NER):** Isolated brand names, competitors, and geographic mentions

4. Sentiment Analysis:

- **VADER** for social media sentiment (optimized for short, informal text)
- **TextBlob** for product reviews
- **Manual Annotation** (10% sample) for model validation

5. Tools Used:

Python (NLTK, spaCy, Scikit-learn), Tableau for visualization

IV. RESULTS

Data Source	Positive (%)	Neutral (%)	Negative (%)	Dominant Themes
Twitter (Brand A)	58	22	20	Pricing, Delivery Time, Support
Amazon Reviews	66	18	16	Quality, Usability, Return Policy
News Headlines	42	38	20	Market Trends, Financial Outlook

Explanation:

- **Twitter data** revealed real-time public perceptions, with spikes in negativity during service delays.
 - **Amazon reviews** showed a strong bias toward positive sentiment, especially for product quality and ease of use.
 - **News headlines** had more neutral language but highlighted shifts in investor sentiment and competitor activity.
- Topic modeling exposed emerging customer concerns about eco-friendly packaging and warranty policies. NER revealed frequent mentions of competitors, suggesting potential market share battles.

V. CONCLUSION

This study demonstrates the potential of integrating text mining and sentiment analysis into strategic market intelligence frameworks. By automating the extraction of qualitative insights from large-scale textual data, businesses can track sentiment shifts, identify competitive threats, and adapt strategies proactively.

The application of sentiment tools tailored for different data sources enhances analytical accuracy. Moreover, combining structured output (from text mining) with visualization enables stakeholders to quickly grasp complex dynamics. Future research can explore multilingual sentiment analysis, sarcasm detection, and integration with real-time dashboards for continuous intelligence.

REFERENCES

1. Mahajan, R. A., Shaikh, N. K., Tikhe, A. B., Vyas, R., & Chavan, S. M. (2022). Hybrid Sea Lion Crow Search Algorithm-based stacked autoencoder for drug sensitivity prediction from cancer cell lines. *International Journal of Swarm Intelligence Research*, 13(1), 21. <https://doi.org/10.4018/IJSIR.304723>
2. Rathod, S. B., Ponnusamy, S., Mahajan, R. A., & Khan, R. A. H. (n.d.). Echoes of tomorrow: Navigating business realities with AI and digital twins. In *Harnessing AI and digital twin technologies in businesses* (Chapter 12). <https://doi.org/10.4018/979-8-3693-3234-4.ch012>
3. A Patel, K., Srinivasulu, A., Jani, K., & Sreenivasulu, G. (2023). Enhancing monkeypox detection through data analytics: a comparative study of machine and deep learning techniques. *Advances in Engineering and Intelligence Systems*, 2(04), 68-80.
4. Shah, M., Bhavsar, N., Patel, K., Gautam, K., & Chauhan, M. (2023, August). Modern Challenges and Limitations in Medical Science Using Capsule Networks: A Comprehensive Review. In *International Conference on Image Processing and Capsule Networks* (pp. 1-25). Singapore: Springer Nature Singapore



5. Shah, M., Vasant, A., & Patel, K. A. (2023, May). Comparative Analysis of Various Machine Learning Algorithms to Detect Cyberbullying on Twitter Dataset. In International Conference on Information, Communication and Computing Technology (pp. 761-787). Singapore: Springer Nature Singapore.
6. Gupta, P. K., Nawaz, M. H., Mishra, S. S., Roy, R., Keshamma, E., Choudhary, S., ... & Sheriff, R. S. (2020). Value Addition on Trend of Tuberculosis Disease in India-The Current Update. *Int J Trop Dis Health*, 41(9), 41-54.
7. Hiremath, L., Kumar, N. S., Gupta, P. K., Srivastava, A. K., Choudhary, S., Suresh, R., & Keshamma, E. (2019). Synthesis, characterization of TiO₂ doped nanofibres and investigation on their antimicrobial property. *J Pure Appl Microbiol*, 13(4), 2129-2140.
8. Gupta, P. K., Lokur, A. V., Kallapur, S. S., Sheriff, R. S., Reddy, A. M., Chayapathy, V., ... & Keshamma, E. (2022). Machine Interaction-Based Computational Tools in Cancer Imaging. *Human-Machine Interaction and IoT Applications for a Smarter World*, 167-186.
9. Gopinandhan, T. N., Keshamma, E., Velmourougane, K., & Raghuramulu, Y. (2006). Coffee husk-a potential source of ochratoxin A contamination.
10. Keshamma, E., Rohini, S., Rao, K. S., Madhusudhan, B., & Udaya Kumar, M. (2008). In planta transformation strategy: an *Agrobacterium tumefaciens*-mediated gene transfer method to overcome recalcitrance in cotton (*Gossypium hirsutum* L.). *J Cotton Sci*, 12, 264-272.
11. Gupta, P. K., Mishra, S. S., Nawaz, M. H., Choudhary, S., Saxena, A., Roy, R., & Keshamma, E. (2020). Value Addition on Trend of Pneumonia Disease in India-The Current Update.
12. Sumanth, K., Subramanya, S., Gupta, P. K., Chayapathy, V., Keshamma, E., Ahmed, F. K., & Murugan, K. (2022). Antifungal and mycotoxin inhibitory activity of micro/nanoemulsions. In *Bio-Based Nanoemulsions for Agri-Food Applications* (pp. 123-135). Elsevier.
13. Hiremath, L., Sruti, O., Aishwarya, B. M., Kala, N. G., & Keshamma, E. (2021). Electrospun nanofibers: Characteristic agents and their applications. In *Nanofibers-Synthesis, Properties and Applications*. IntechOpen.
14. Hussain, M. M. A. *Business Analytics: The Key to Smarter, Faster, and Better Decisions*.
15. Hussain, M. A. (2013). Impact of visual merchandising on consumer buying behaviour at big bazaar. *International Journal of retail and distribution management*, 3(2).
16. Hussain, M. A., Gupta, R., Kushwaha, A., Samanta, P., Khulbe, M., & Ahmad, V. (2024, June). Transforming technology for online marketing with focus on artificial intelligence: a qualitative approach. In 2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT) (pp. 1-5). IEEE.
17. Das, A., Shobha, N., Natesh, M., & Tiwary, G. (2024). An Enhanced Hybrid Deep Learning Model to Enhance Network Intrusion Detection Capabilities for Cybersecurity. *Journal of Machine and Computing*, 4(2), 472.
18. Gowda, S. K., Murthy, S. N., Hiremath, J. S., Subramanya, S. L. B., Hiremath, S. S., & Hiremath, M. S. (2023). Activity recognition based on spatio-temporal features with transfer learning. *Int J Artif Intell* ISSN, 2252(8938), 2103.
19. Shanthala, K., Chandrakala, B. M., & Shobha, N. (2023, November). Automated Diagnosis of brain tumor classification and segmentation of MRI Images. In 2023 International Conference on the Confluence of Advancements in Robotics, Vision and Interdisciplinary Technology Management (IC-RVITM) (pp. 1-7). IEEE.
20. Nagar, H., & Menaria, A. K. Compositions of the Generalized Operator $(G\rho, \eta, \gamma, \omega; a \Psi)(x)$ and their Application.
21. NAGAR, H., & MENARIA, A. K. (2012). Applications of Fractional Hamilton Equations within Caputo Derivatives. *Journal of Computer and Mathematical Sciences* Vol, 3(3), 248-421.
22. Nagar, H., & Menaria, A. K. On Generalized Function $G\rho, \eta, \gamma [a, z]$ And It's Fractional Calculus.
23. Suma, V., & Nair, T. G. (2008, October). Enhanced approaches in defect detection and prevention strategies in small and medium scale industries. In 2008 The Third International Conference on Software Engineering Advances (pp. 389-393). IEEE.
24. Rashmi, K. S., Suma, V., & Vaidehi, M. (2012). Enhanced load balancing approach to avoid deadlocks in cloud. arXiv preprint arXiv:1209.6470.
25. Nair, T. G., & Suma, V. (2010). The pattern of software defects spanning across size complexity. *International Journal of Software Engineering*, 3(2), 53-70.
26. Rao, Jawahar J., and V. Suma. "Effect of Scope Creep in Software Projects—Its Bearing on Critical SuccessFactors." *International Journal of Computer Applications* 975 (2014): 8887.
27. Suma, V. (2020). Automatic spotting of sceptical activity with visualization using elastic cluster for network traffic in educational campus. *Journal of Ubiquitous Computing and Communication Technologies*, 2, 88-97.
28. Nair, TR Gopalakrishnan, and V. Suma. "A paradigm for metric based inspection process for enhancing defect management." *ACM SIGSOFT Software Engineering Notes* 35, no. 3 (2010): 1.
29. Polamarasetti, S. (2021). Evaluating the Effectiveness of Prompt Engineering in Salesforce Prompt Studio. *International Journal of Emerging Trends in Computer Science and Information Technology*, 2(3), 96-103.



30. Rajoria, N. V., & Menaria, A. K. Numerical Approach of Fractional Integral Operators on Heat Flux and Temperature Distribution in Solid.
31. Polamarasetti, S. (2022). Using Machine Learning for Intelligent Case Routing in Salesforce Service Cloud. *International Journal of AI, BigData, Computational and Management Studies*, 3(1), 109-113.
32. Polamarasetti, S. (2021). Enhancing CRM Accuracy Using Large Language Models (LLMs) in Salesforce Einstein GPT. *International Journal of Emerging Trends in Computer Science and Information Technology*, 2(4), 81-85.
33. Sahoo, S. C., Sil, A., Solanki, R. T., & Dutta, A. (2023). Fire Performance and Technological Properties of Plywood Prepared by with PMUF Adhesive Modified with Organic Phosphate. *J. Chem. Heal. Risks*, 13, 2627-2637.
34. Sil, A. (2016). Study on Bamboo Composites as Components of Housing System for Disaster Prone Areas. *International Journal of Civil Engineering (IJCE)*, 5(3), 11-18.
35. Sahoo, S. C., Sil, A., & Solanki, R. T. (2020). Effect of adhesive performance of liquid urea formaldehyde (UF) resin when used by mixing with solid UF resin for manufacturing of wood based panels. *Int. J. Sci. Res. Publ*, 10, 10065.
36. Sil, A. (2022). Bamboo—A green construction material for housing towards sustainable economic growth. *Int. J. Civ. Eng. Technol*, 13, 1-9.
37. Sahoo, S. C., Sil, A., Thanigai, K., & Pandey, C. N. (2011). Use of silicone based coating for protection of wood materials and bamboo composites from weathering and UV degradation. *Journal of the Indian Academy of Wood Science*, 8(2), 143-147.